Plant Molecular Biology **22**: 301–312, 1993.

Identification and expression of a cDNA from *Daucus carota* encoding a bifunctional aspartokinase-homoserine dehydrogenase[1]

Jane M. Weisemann and Benjamin F. Matthews*
*United States Department of Agriculture, Agricultural Research Service, Plant Molecular Biology Laboratory, Beltsville, MD 20705-2350, USA (*author for correspondence)*

Received 16 November 1992; accepted in revised form 24 February 1993

Key words: aspartokinase, homoserine dehydrogenase, cDNA, carrot, chloroplast, gene

Abstract

Aspartokinase (EC 2.7.2.4) and homoserine dehydrogenase (EC 1.1.1.3) catalyze steps in the pathway for the synthesis of lysine, threonine, and methionine from aspartate. Homoserine dehydrogenase was purified from carrot (*Daucus carota* L.) cell cultures and portions of it were subjected to amino acid sequencing. Oligonucleotides deduced from the amino acid sequences were used as primers in a polymerase chain reaction to amplify a DNA fragment using DNA derived from carrot cell culture mRNA as template. The amplification product was radiolabelled and used as a probe to identify cDNA clones from libraries derived from carrot cell culture and root RNA. Two overlapping clones were isolated. Together the cDNA clones delineate a 3089 bp long sequence encompassing an open reading frame encoding 921 amino acids, including the mature protein and a long chloroplast transit peptide. The deduced amino acid sequence has high homology with the *Escherichia coli* proteins aspartokinase I-homoserine dehydrogenase I and aspartokinase II-homoserine dehydrogenase II. Like the *E. coli* genes the isolated carrot cDNA appears to encode a bifunctional aspartokinase-homoserine dehydrogenase enzyme.

Abbreviations: AK, aspartokinase; HSDH, homoserine dehydrogenase; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate

Introduction

In plants, the essential amino acids lysine, threonine, methionine and isoleucine are synthesized from aspartate. These amino acids are required in the diets of non-ruminant animals, thus, there is much interest in increasing the quantity of these amino acids in food sources [1]. The first reaction in the pathway, the conversion of aspartate into β-aspartyl phosphate, is catalyzed by the enzyme

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number L11529.

[1] The mention of vendor or product does not imply that they are endorsed or recommended by the U.S. Department of Agriculture over vendors of similar products not mentioned.

aspartokinase (AK; EC 2.7.2.4). In the second reaction of the pathway, β-aspartyl phosphate is converted to aspartate semialdehyde. At this point the pathway branches: one branch leads to the formation of lysine and the other to the formation of threonine and isoleucine or methionine. The first reaction in this second branch is the reduction of β-aspartyl phosphate to homoserine by the enzyme homoserine dehydrogenase (HSDH; EC 1.1.1.3). These biosynthetic pathways in plants are similar to pathways found in bacteria [4, 5]. Many of the bacterial genes that code for the enzymes of the aspartate pathway have been cloned and sequenced [5]. In contrast the only plant gene coding for an enzyme of the aspartate pathway which has been isolated is that for dihydrodipicolinate synthase, which catalyzes the first reaction specific to lysine synthesis [8].

In most of the plant species studied multiple forms of AK and HSDH have been identified *in vitro* [1]. These forms are distinguished by their sensitivity to feedback inhibition (in particular by threonine and lysine) and by their molecular weight and subunit composition. From carrot (*Daucus carota* L.) suspension cultures two forms of HSDH have been identified *in vitro*: one sensitive and one insensitive to threonine inhibition [13, 21]. These two forms can interconvert. Three forms of AK have been isolated from carrot: form I is strongly inhibited by lysine, form II is strongly inhibited by threonine, and form III is partially inhibited by both [24]. The relationship between these three forms has not yet been defined.

The isolation of clones of these genes from plants should enable us to answer questions about the relationship between various forms of the enzymes, the number of genes involved, and the regulation of the pathway. In *E. coli* there are three genes coding for AK and HSDH [4, 5]. One, *lysC*, codes for a lysine-sensitive AK and is regulated by lysine. The other two genes code for bifunctional AK-HSDH proteins. Expression of the *thrA* gene is repressed by threonine and isoleucine and the enzymatic activity of AKI-HSDHI is inhibited by threonine. Expression of the *metL* gene is repressed by methionine, but the protein AKII-HSDHII is not responsive to end-product inhibition [4, 5]. It is not known if the multiple enzyme forms in plants are encoded by separate genes or if these genes are subject to transcriptional or translational regulation. Knowledge gained from the study of the aspartate pathway genes may eventually allow the engineering of the pathway to alter the amino acid pool composition of plants used as protein sources.

Our previous biochemical studies of carrot AK and HSDH indicated that these activities resided on the same protein molecule as a bifunctional enzyme [24]. This paper describes the isolation of a carrot cDNA encoding a bifunctional AK-HSDH. In fact, this is the first report of a gene encoding a bifunctional AK-HSDH protein outside of the Enterobacteriaceae.

Materials and methods

General DNA cloning methods

Plasmid DNA preparation, ligations, restriction enzyme digestions, Southern blots, and northern blots were done according to standard procedures [12]. Plant genomic DNA was extracted by the method of Keim *et al.* [10]. Enzymes were obtained from either Bethesda Research Laboratories or Boehringer Mannheim. PCR reagents were obtained from Perkin-Elmer Cetus. Oligonucleotides used for PCR were obtained from Synthecell.

DNA probes were labelled with [α-^{32}P] dCTP (DuPont-New England Nuclear) using the random oligo priming method [7].

DNA sequence determination was done by the dideoxy chain termination method using modified T7 DNA polymerase (Sequenase 2.0 from United States Biochemical Corp.) [20].

Construction of a carrot cDNA library

Total RNA was extracted from two-month old carrot roots by the method of Chirgwin *et al.* [3].

Poly(A)$^+$ RNA was isolated from total RNA on columns of oligo(dT) cellulose [12]. cDNA was made from 3 μg poly (A)$^+$ RNA using the cDNA Synthesis System Plus from Amersham. *Eco* RI adapters were added (Promega 'Riboclone *Eco* RI Adaptor Ligation System') and the cDNA was ligated from λgt11 arms (Promega) and packaged using Stratagene 'Gigapack II Plus' packaging extracts. The phage were plated and screened following methods described by Maniatis *et al.* [12]. Approximately 240000 plaques were screened using a radiolabelled PCR product as the probe. Plaque lifts were done in duplicate onto nitrocellulose filters. Filters were hybridized in 50% formamide, 5 × Denhardt's solution, 5 × SSPE (1 × SSPE = 0.15 M NaCl, 0.01 M sodium phosphate pH 7.4, 0.001 M EDTA), 0.1% SDS and 100 μg/ml denatured salmon sperm DNA at 42 °C for 18 h. The final wash was with 0.1 × SSC (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate), 0.1% SDS at 45 °C for one set of filters and 60 °C for the second set.

A second cDNA library was constructed in the same manner, except that poly (A)$^+$ RNA from carrot cell culture was used and first-strand synthesis was primed with an oligonucleotide corresponding to a tryptic peptide from purified carrot HSDH [24] (HSDH97: 5′-GCT/CTCA/GTAA/GAAA/GTAA/GTGNGTA/GTA-3′).

DNA and RNA blot analysis

Genomic DNA was isolated from carrot leaves [10], digested with restriction enzymes according to the manufacturer's instructions, and 10 μg from each digestion was separated by electrophoresis on a 0.7% agarose gel and transferred to nitrocellulose by capillary action [22]. The blot was probed with the full-length AK-HSDH clone. Hybridization was in 25% formamide, 5 × SSPE, 5 × Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42 °C for 23 hr. The final wash was in 2 × SSC, 0.1% SDS at 42 °C.

Total RNA was isolated from carrot cell suspension cultures at different times after inoculation and from 2-month old leaves and roots. Total RNA (10 μg) from each sample was separated on an agarose-formaldehyde-formamide gel and transferred to a nitrocellulose membrane. Gels were stained with ethidium bromide prior to transfer to confirm that equal quantities of RNA were loaded in each lane and that transfer to nitrocellulose was complete. The blot was hybridized with a ^{32}P-labelled 300 bp *Eco* RI-*Hin*c II fragment from the 5′ end of the clone. Hybridization was in 50% formamide, 5 × SSPE, 5 × Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42 °C for 18 h and washed as described above at 42 °C.

PCR amplification of a carrot cDNA encoding AK-HSDH

Protein with HSDH activity was purified from carrot cell suspension cultures [13, 24]. The purified protein was subjected to proteolysis, the polypeptides were separated by HPLC, and the amino acid sequences of four polypeptides were determined [24]. The locations of these polypeptides are presented in Fig. 1. Based on the amino acid sequence of two of these polypeptides, the following degenerate oligonucleotides were synthesized:

Oligo HSDH 76: 5′-TAT/C CAA/G GAA/G GCN TGG GAA/G ATG-3′

(Peptide 76: NH_2-Y Q E A W E M)

Oligo HSDH 97: 5′-GC T/CTC A/GTA A/GAA A/GTA A/GTG NGT A/GTA-3′

(Peptide 97: A E Y F Y H T Y-NH_2).

These oligonucleotides were used as primers for PCR amplification of carrot cDNA. Both double-stranded cDNA made from carrot root poly(A)$^+$ RNA and first-strand cDNA made from carrot cell culture poly(A)$^+$ RNA served as templates. When the amplification was carried out at an annealing temperature of 52 °C, using either template, we obtained predominantly a 1100 bp product (data not shown). The 579 bp *Eco* RI-*Hin*d III fragment from the PCR product was subcloned into vectors M13mp18 and mp19 and the DNA sequence was determined.

```
1                                                                    60
GTCGCTCTCTTCCGCCATATCTCCGTCATCTTATGCCGCCATCGCCGCCGCTTACTCTGC
 S  L  S  S  A  I  S  P  S  S  Y  A  A  I  A  A  A  Y  S  A

61                                                                   120
ACGGACTCCCATTTTCAACAAGAAAAAGACGGCGGCGGTTCTCTCTCCCCTCTCTCTTTT
 R  T  P  I  F  N  K  K  K  T  A  A  V  L  S  P  L  S  L  F

121                                                                  180
CCATCAGTCTCCCTCTCTCTCCAAAACAGGTATATTTTTGCATCGGGGAAGAAAGGAGTC
 H  Q  S  P  S  L  S  K  T  G  I  F  L  H  R  G  R  K  E  S

181                                                                  240
GTCGTCGAAGTTTTACATTGCTGCTTCCGTTACAACTGCAGTTCCTTCTCTCGATGACTC
 S  S  K  F  Y  I  A  A  S  V  T  T  A  V  P  S  L  D  D  S

241                                                                  300
CGTTGAGAAGGTTCACCTTCCCAGGGGTGCTATGTGGTCTATTCATAAATTTGGAGGCAC
 V  E  K  V  H  L  P  R  G  A  M  W  S  I  H  K  F  G  G  T

301                                                                  360
CTGTGTGGGAAGCTCTGAAAGGATCCGAAATGTTGCAGAGATAGTTGTGGAGGATGATTC
 C  V  G  S  S  E  R  I  R  N  V  A  E  I  V  V  E  D  D  S

361                                                                  420
TGAAAGAAAGCTAGTTGTAGTCTCTGCAATGTCAAAGGTCACAGATATGATGTATGATCT
 E  R  K  L  V  V  V  S  A  M  S  K  V  T  D  M  M  Y  D  L

421                                                                  480
AATTTACAAGGCGCAGTCACGGGATGATTCTTATGAATCTGCGCTCGATGCTGTTATGGA
 I  Y  K  A  Q  S  R  D  D  S  Y  E  S  A  L  D  A  V  M  E

481                                                                  540
AAAGCACAAGTTAACAGCATTTGATCTCCTTGATGAAGATGACCTTGCTAGATTTTTAAC
 K  H  K  L  T  A  F  D  L  L  D  E  D  D  L  A  R  F  L  T

541                                                                  600
TAGGCTGCAACATGATGTTATTACCCTCAAAGCAATGCTTCGTGCAATATACATAGCTGG
 R  L  Q  H  D  V  I  T  L  K  A  M  L  R  A  I  Y  I  A  G

601                                                                  660
TCATGCCACCGAATCTTTTTCGGATTTTGTTGTGGGACATGGAGAGCTATGGTCAGCTCA
 H  A  T  E  S  F  S  D  F  V  V  G  H  G  E  L  W  S  A  Q

661                                                                  720
GCTGTTGTCATTTGTAATAAGAAAGAATGGGGGTGACTGTAATTGGATGGACACACGAGA
 L  L  S  F  V  I  R  K  N  G  G  D  C  N  W  M  D  T  R  D

721                                                                  780
TGTTCTTGTTGTAAATCCTGCTGGATCTAATCAAGTCGATCCTGATTATTTGGAATCTGA
 V  L  V  V  N  P  A  G  S  N  Q  V  D  P  D  Y  L  E  S  E

781                                                                  840
GAAGAGACTTGAGAAATGGTTCTCCAGCAACCAGTGTCAGACAATTGTTGCGACAGGTTT
 K  R  L  E  K  W  F  S  S  N  Q  C  Q  T  I  V  A  T  G  F

841                                                                  900
TATAGCTAGCACGCCTCAAAATATACCTACAACTTTGAAAAGAGACGGAAGTGACTTTTC
 I  A  S  T  P  Q  N  I  P  T  T  L  K  R  D  G  S  D  F  S

901                                                                  960
TGCCGCTATAATGGGTGCTTTATTAAGGGCTGGTCAAGTCACGATTTGGACTGATGTTAA
 A  A  I  M  G  A  L  L  R  A  G  Q  V  T  I  W  T  D  V  N

961                                                                 1020
TGGTGTATATAGTGCAGATCCTCGAAAAGTTAGTGAGGCTGTGGTATTAAAGACATTGTC
 G  V  Y  S  A  D  P  R  K  V  S  E  A  V  V  L  K  T  L  S
                                                 Pep 76 T  L  S

1021                                                                1080
TTATCAAGAAGCCTGGGAGATGTCATATTTTGGGGCTAATGTGTTACATCCCCGTACTAT
 Y  Q  E  A  W  E  M  S  Y  F  G  A  N  V  L  H  P  R  T  I
 Y  Q  E  A  W  E  M  S  Y  F  G  A  N  V  L  H  P  R

1081                                                                1140
CAATCCTGTGATGCGATATGACATTCCAATTGTAATAAGAAATATATTCAACCTATCTGC
 N  P  V  M  R  Y  D  I  P  I  V  I  R  N  I  F  N  L  S  A

1141                                                                1200
TCCGGGAACAATGATATGCCGAGAATCTGTAGGCGAAACTGAAGATGGGTTAAAATTGGA
 P  G  T  M  I  C  R  E  S  V  G  E  T  E  D  G  L  K  L  E

1201                                                                1260
ATCTCATGTCAAAGGATTTGCTACTATTGATAATCTGGCGCTCATTAATGTTGAAGGAAC
 S  H  V  K  G  F  A  T  I  D  N  L  A  L  I  N  V  E  G  T
      Pep 90 G  F  A  T  I  D  N  L  A  L  I  N  V  E  G  T
```

```
1261                                                                1261
TGGAATGGCTGGTGTTCCTGGTACAGCTACTGCAATTTTTGGTGCTGTCAAGGATGTGGG
 G  M  A  G  V  P  G  T  A  T  A  I  F  G  A  V  K  D  V  G
 G  M  A  G  V  P  G  T  A  T  A  I  F [S] A  V (K)

1321                                                                1380
AGCTAATGTTATAATGATATCTCAGGCTAGCAGTGAGCATTCTATTTGCTTTGCTGTGCC
 A  N  V  I  M  I  S  Q  A  S  S  E  H  S  I  C  F  A  V  P

1381                                                                1440
TGAGAGTGAAGTGAAAGCTGTTGCTAAAGCTTTGGAGGCCAGATTTCGTCAAGCTTTAGA
 E  S  E  V  K  A  V  A  K  A  L  E  A  R  F  R  Q  A  L  D

1441                                                                1500
TGCTGGTCGTCTTTCCCAGGTTGCTAATAATCCAAACTGTAGCATCTTGGCAACAGTTGG
 A  G  R  L  S  Q  V  A  N  N  P  N  C  S  I  L  A  T  V  G

1501                                                                1560
CCAAAAGATGGCAAGTACTCCTGGCGTGAGTGCTACACTTTTCAATGCGCTTGCAAAGGC
 Q  K  M  A  S  T  P  G  V  S  A  T  L  F  N  A  L  A  K  A

1561                                                                1620
CAATATAAACGTTCGTGCTATAGCCCAGGGCTGTACAGAGTATAATATCACTGTAGTTCT
 N  I  N  V  R  A  I  A  Q  G  C  T  E  Y  N  I  T  V  V  L

1621                                                                1680
CAGTCGAGAAGATTGTGTGAGGGCTTTGAAAGCTGTCCATTCAAGATTTTATCTGTCGAG
 S  R  E  D  C  V  R  A  L  K  A  V  H  S  R  F  Y  L  S  R

1681                                                                1740
AACCACAATAGCAGTGGGTATTGTCGGACCTGGATTAATCGGAGCTACTTTACTTGACCA
 T  T  I  A  V  G  I  V  G  P  G  L  I  G  A  T  L  L  D  Q

1741                                                                1800
GCTCAGGGATCAGGCAGCAATCCTCAAGGAAAATTCTAAAATTGATTTGCGTGTTATGGG
 L  R  D  Q  A  A  I  L  K  E  N  S  K  I  D  L  R  V  M  G

1801                                                                1860
TATCACCGGATCGAGAACAATGCTTCTGAGCGAAACGGGAATCGATTTAAGTAGATGGAG
 I  T  G  S  R  T  M  L  L  S  E  T  G  I  D  L  S  R  W  R

1861                                                                1920
AGAAGTCCAAAAAGAGAAAGGGCAAACAGCTGGCCTAGAAAAATTTGTACAACATGTGCG
 E  V  Q  K  E  K  G  Q  T  A  G  L  E  K  F  V  Q  H  V  R

1921                                                                1980
TGGAAATCATTTTATTCCAAGCACTGTTATAGTAGATTGTACAGCAGACTCTGAAGTGGC
 G  N  H  F  I  P  S  T  V  I  V  D  C  T  A  D  S  E  V  A

1981                                                                2040
AAGTCACTACCATGACTGGTTGTGTAGGGGAATTCACGTCATTACCCCAAATAAGAAGGC
 S  H  Y  H  D  W  L  C  R  G  I  H  V  I  T  P  N  K  K  A

2041                                                                2100
AAATTCAGGACCCCTTGATCAGTATTTGAAGTTGAGAGCTCTCCAGCGGCGATCCTATAC
 N  S  G  P  L  D  Q  Y  L  K  L  R  A  L  Q  R  R  S  Y  T
                                              Pep 97 S  Y  T

2101                                                                2160
ACACTATTTCTATGAAGCTACTGTTGTTGCTGGTCTCCCGATCATAACCACTTTGCAGGG
 H  Y  F  Y  E  A  T  V  V  A  G  L  P  I  I  T  T  L  Q  G
 H  Y  F  Y  E  A  T  V  V  A  G  L  P  I  I  T  T  L  Q  G

2161                                                                2220
ACTTCTTGAAACCGGGGACAAGATATTGCGAATTGAAGGCATTTTCAGTGGGACTCTTAG
 L  L  E  T  G  D  K  I  L  R  I  E  G  I  F  S  G  T  L  S
 L  L  E  T  G  D
2221                                                                2280
TTACATATTCAACAACTTTAAGAGTACAACACCTTTTAGTGAAGTGGTAAGTGAGGCAAA
 Y  I  F  N  N  F  K  S  T  T  P  F  S  E  V  V  S  E  A  K

2281                                                                2340
AGCGGCAGGGTATACTGAACCAGATCCAAGGGATGATCTAGCCGGAACTGATGTTGCTAG
 A  A  G  Y  T  E  P  D  P  R  D  D  L  A  G  T  D  V  A  R

2341                                                                2400
AAAGGTAATAATTCTTGCTAGAGGATCTGGATTAAAGCTCGAACTGTCTGATATCCCTGT
 K  V  I  I  L  A  R  G  S  G  L  K  L  E  L  S  D  I  P  V

2401                                                                2460
ACAGAGCCTTGTTCCAGAACCACTAAGGGGTATTGCGTCAGCCGAAGAATTTCTGCTACA
 Q  S  L  V  P  E  P  L  R  G  I  A  S  A  E  E  F  L  L  Q
```

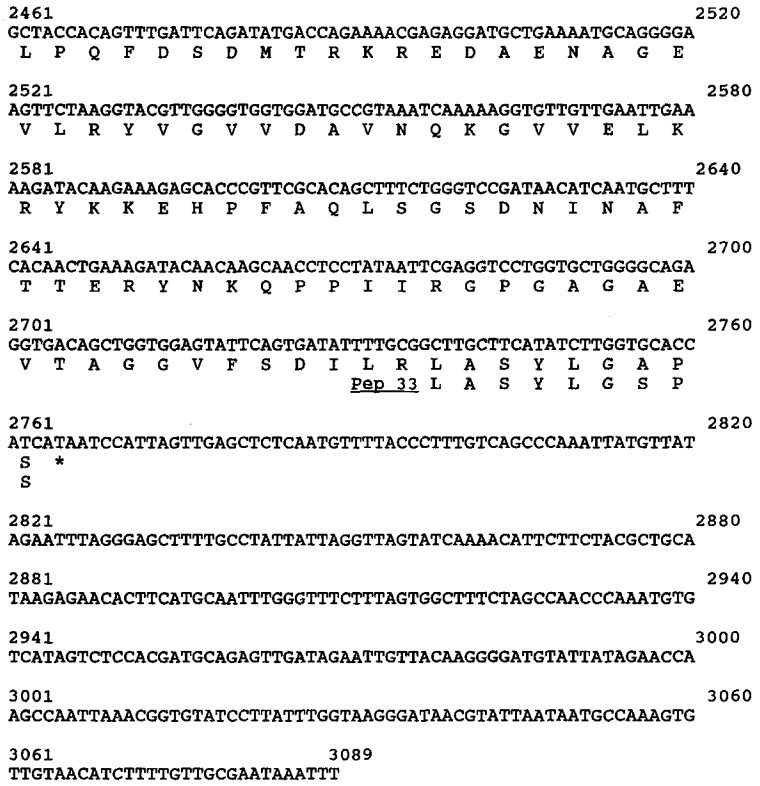


Fig. 1. DNA sequence of carrot of the AK-HSDH cDNA clone and the deduced amino acid sequence. Also shown are the positions of the peptides whose amino acid sequences were determined.

Isolation of λgt11 carrot cDNA clones for HSDH

The 1100 bp PCR-amplified DNA fragment was used to screen a carrot cDNA library. The library was constructed in the vector λgt11 and contained cDNA from carrot root poly(A)$^+$ RNA. From approximately 250000 plaques four clones were obtained. All of the clones displayed similar restriction digestion patterns and appeared to differ from one another only in length. The complete DNA sequence of the longest clone, HSDH1, was determined. The clone was 2079 bp long and had one long open reading frame starting at bp 1 and reading through bp 1755 for a total of 585 codons.

Clone HSDH1 had no apparent translation start site. RNA blot analysis using the HSDH PCR product as a probe identified a 3000 nucleotide transcript as the main hybridizing band (not shown). The open reading frame in the clone was not sufficient to code for a protein of 85 ± 5 kDa (the subunit size of the purified carrot enzyme [13]). These facts indicated that clone HSDH1 was lacking ca. 1000 bp at the 5′ end.

Isolation of λgt11 carrot cDNA clones for the 5′ end of the HSDH gene

In order to enhance the probability of finding HSDH clones with a complete 5′ end, we constructed a second cDNA library using oligonucleotide HSDH97 to prime first-strand synthesis. Poly (A)$^+$ RNA was isolated from carrot cell suspension cultures. The cDNA was cloned as above into the vector λgt11. 350000 plaques were screened and 14 positive clones were analyzed. Three of these contained inserts that extended

beyond the 5′ end of HSDH1. The longest of these clones, HSDH97-12, was 1009 bp longer at the 5′ end than HSDH1.

DNA sequence analysis and comparison to other genes coding for HSDH and AK

The DNA sequence was determined for the total length of clones HSDH1 and HSDH97-12 and for parts of other clones. Figure 1 shows the 3089 bp DNA sequence of the combined HSDH1 and HSDH97-12 clones and the deduced amino acid sequence for the open reading frame from bp 2 through bp 2764. The sequences of four tryptic peptides from the purified carrot HSDH had been determined [24]. These were compared to the deduced amino acid sequence from the cDNA clone. Only one amino acid residue out of 92 did not match ([S] in peptide 90) and this residue was ambiguous in the peptide sequence.

The amino-terminal sequence for carrot HSDH could not be determined from the purified protein. The sequence of the cDNA clone shows a possible translation start site (ATG) at bp 272. If translation were to start at this site, the resulting 831 amino acid protein would have a molecular weight of 90679 Da, which is in the expected size range.

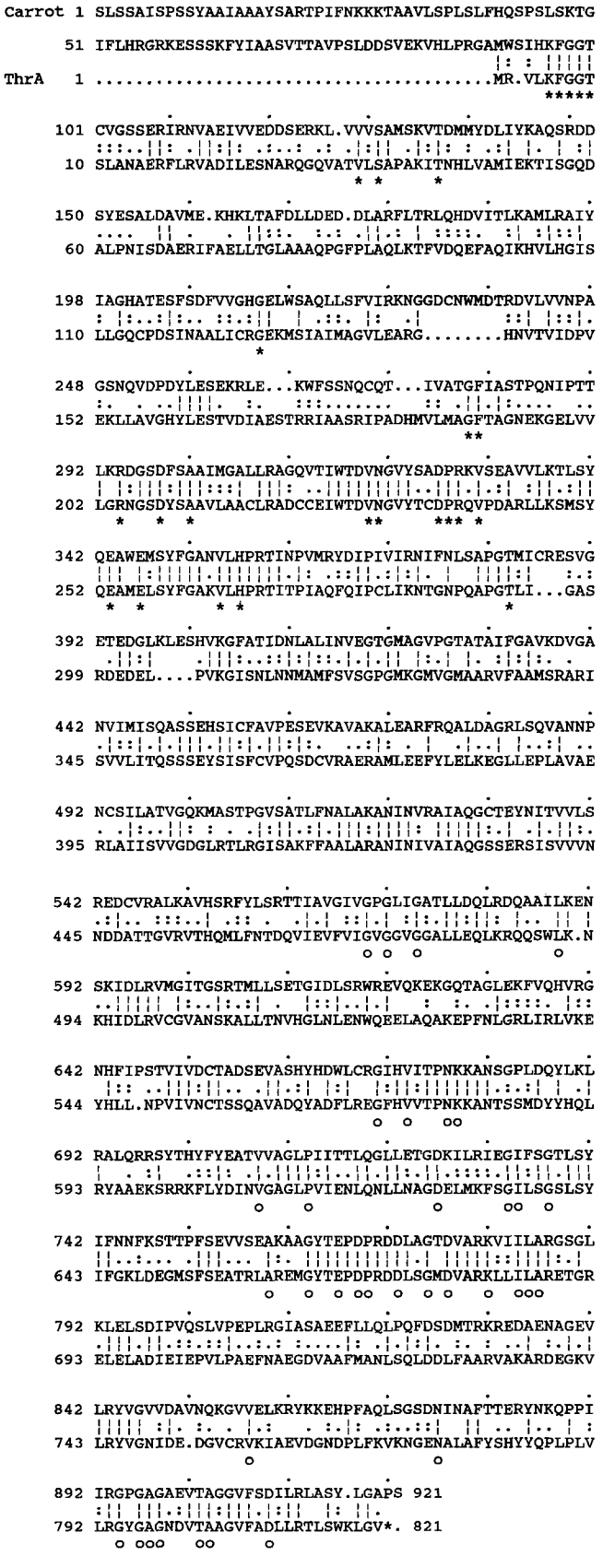
The deduced amino acid sequence of the carrot HSDH clone was compared to bacterial and fungal HSDH and AK proteins for which there is sequence information available (Table 1). The strongest homology was to the bifunctional *E. coli* AKI-HSDHI encoded by the *thrA* gene (Fig. 2). *E. coli* AKI-HSDHI is 820 amino acids long and can be divided into three general regions. Ca. 248 amino acids at the amino-terminus constitute the AK functional domain; about 359 residues (471–820) at the carboxy-terminus constitute the HSDH domain, and residues from about 249 to 470 make up a central interface domain. AKII-HSDHII has a similar structure [4, 5]. Proteins with only AK function (*E. coli* AKIII, *Bacillus subtilis* AKII, yeast AK) have homology to amino acids 1 through about 500 of AKI-HSDHI. Proteins with only HSDH function (*B. subtilis* HSDH, *Corynebacterium glutamicum* HSDH) have homology beginning at about amino acid 500 and extending through the carboxy terminus of AKI-HSDHI. There is homology between the carrot HSDH and *E. coli* AKI-HSDHI along the full length of the proteins, although it is strongest in the HSDH domain and the interface domain. The homology with AKII-HSDHII is at a simi-

Table 1. Homology between carrot AK-HSDH and AK and HSDH proteins from other organisms. Comparisons of the carrot AK-HSDK to the three enzymes from *E. coli*, to the AK of *Bacillus subtilis*, and to the HSDH from *Corynebacterium glutamicum* were made using the GAP program of the University of Wisconsin Genetics Computer Group. The comparisons to the *B. subtilis* HSDH and *Saccharomyces cerevisiae* AK were based on published alignments between these proteins and the *E. coli* proteins [16, 17].

	Identity (%)	Homology[1] (%)	Total identity + homology (%)
E. coli AKI-HSDHI	38	10	48
Amino acids 91–385[2]	30	9	39
Amino acids 386–567	40	9	49
Amino acids 568–921	40	10	50
E. coli AKII-HSDHII	33	9	42
Amino acids 91–385	32	7	39
Amino acids 386–567	19	9	28
Amino acids 568–921	38	10	48
E. coli AKIII (Amino acids 91–385)	28	12	40
B. subtilis AKII (Amino acids 91–558)	33	9	42
B. subtilis HSDH (Amino acids 91–921)	22	8	30
C. glutamicum HSDH (Amino acids 547–921)	25	11	36
S. cerevisiae AK (Amino acids 91–463)	19	8	27

[1] Amino acid replacements: Ile-Val-Leu, Ser-Thr, Lys-Arg, Phe-Tyr, Glu-Asp.

[2] Amino acid residue numbers refer to the amino acids of the carrot protein as shown in Fig. 2.


Carrot 1 SLSSAISPSSYAAIAAAYSARTPIFNKKKTAAVLSPLSLFHQSPSLSKTG 50
51 IFLHRGRKESSSKFYIAASVTTAVPSLDDSVEKVHLPRGAMWSIHKFGGT 100
ThrA 1 .....MR.VLKFGGT 9
101 CVGSSERIRNVAEIVVEDDSERKL.VVVSAMSKVTDMMYDLIYKAQSRDD 149
10 SLANAERFLRVADILESNARQGQVATVLSAPAKITNHLVAMIEKTISGQD 59
150 SYESALDAVME.KHKLTAFDLLDED.DLARFLTRLQHDVITLKAMLRAIY 197
60 ALPNISDAERIFAELLTGLAAAQPGFPLAQLKTFVDQEFAQIKHVLHGIS 109
198 IAGHATESFSDFVVGHGELWSAQLLSFVIRKNGGDCNWMDTRDVLVVNPA 247
110 LLGQCPDSINAALICRGEKMSIAIMAGVLEARG.....HNVTVIDPV 151
248 GSNQVDPDYLESEKRLE...KWFSSNQCQT...IVATGFIASTPQNIPTT 291
152 EKLLAVGHYLESTVDIAESTRRIAASRIPADHMVLMAGFTAGNEKGELVV 201
292 LKRDGSDFSAAIMGALLRAGQVTIWTDVNGVYSADPRKVSEAVVLKTLSY 341
202 LGRNGSDYSAAVLAACLRADCCEIWTDVNGVYTCDPRQVPDARLLKSMSY 251
342 QEAWEMSYFGANVLHPRTINPVMRYDIPIVIRNIFNLSAPGTMICRESVG 391
252 QEAMELSYFGAKVLHPRTITPIAQFQIPCLIKNTGNPQAPGTLI...GAS 298
392 ETEDGLKLESHVKGFATIDNLALINVEGTGMAGVPGTATAIFGAVKDVGA 441
299 RDEDEL....PVKGISNLNNMAMFSVSGPGMKGMVGMAARVFAAMSRARI 344
442 NVIMISQASSEHSICFAVPESEVKAVAKALEARFRQALDAGRLSQVANNP 491
345 SVVLITQSSSEYSISFCVPQSDCVRAERAMLEEFYLELKEGLLEPLAVAE 394
492 NCSILATVGQKMASTPGVSATLFNALAKANINVRAIAQGCTEYNITVVLS 541
395 RLAIISVVGDGLRTLRGISAKFFAALARANINIVAIAQGSSERSISVVVN 444
542 REDCVRALKAVHSRFYLSRTTIAVGIVGPGLIGATLLDQLRDQAAILKEN 591
445 NDDATTGVRVTHQMLFNTDQVIEVFVIGVGGVGGALLEQLKRQQSWLK.N 493
592 SKIDLRVMGITGSRTMLLSETGIDLSRWREVQKEKGQTAGLEKFVQHVRG 641
494 KHIDLRVCGVANSKALLTNVHGLNLENWQEELAQAKEPFNLGRLIRLVKE 543
642 NHFIPSTVIVDCTADSEVASHYHDWLCRGIHVITPNKKANSGPLDQYLKL 691
544 YHLL.NPVIVNCTSSQAVADQYADFLREGFHVVTPNKKANTSSMDYYHQL 592
692 RALQRRSYTHYFYEATVVAGLPIITTLQGLLETGDKILRIEGIFSGTLSY 741
593 RYAAEKSRRKFLYDINVGAGLPVIENLQNLLNAGDELMKFSGILSGSLSY 642
742 IFNNFKSTTPFSEVVSEAKAAGYTEPDPRDDLAGTDVARKVIILARGSGL 791
643 IFGKLDEGMSFSEATRLAREMGYTEPDPRDDLSGMDVARKLLILARETGR 692
792 KLELSDIPVQSLVPEPLRGIASAEEFLLQLPQFDSDMTRKREDAENAGEV 841
693 ELELADIEIEPVLPAEFNAEGDVAAFMANLSQLDDLFAARVAKARDEGKV 742
842 LRYVGVVDAVNQKGVVELKRYKKEHPFAQLSGSDNINAFTTERYNKQPPI 891
743 LRYVGNIDE.DGVCRVKIAEVDGNDPLFKVKNGENALAFYSHYYQPLPLV 791
892 IRGPGAGAEVTAGGVFSDILRLASY.LGAPS 921
792 LRGYGAGNDVTAAGVFADLLRTLSWKLGV*. 821

lar level, except in the interface region, where it is reduced.

All AK amino acid sequences analyzed have the sequence KFGGT near the amino-terminus. This includes the yeast and *B. subtilis* AKs, *E. coli* AKIII, and *E. coli* AKI-HSDHI. (*E. coli* AKII-HSDHII has the variation KFGGS). The deduced amino acid sequence of the carrot HSDH clone also has this sequence. Another highly conserved region contains the Gly-X-Gly-X-X-Gly motif which is presumed to be part of the NADH/NADPH binding domain of HSDH [16]. This is found in the carrot sequence at amino acids 570–575 as numbered in Fig. 2.

DNA and RNA blot analysis

All the carrot cDNA clones isolated appeared to represent the same gene. However, it was possible that there were additional genes for AK or HSDH. To test for the existence of other related mRNAs we probed total carrot RNA with the AK-HSDH clone (Fig. 3). Northern blots showed predominantly a message of about 3000 nucleotides. Very minor bands were also seen at about 1500 and 1250 nucleotides. It has not been determined if these are breakdown products of the larger RNA or if they represent second messages. Although Fig. 3 shows a blot probed with a 300 bp fragment from the 5′ end of the clone, a similar pattern was seen when the full-length cDNA was used. Little change was seen in signal levels from samples taken at intervals from one to seven days of cell suspension culture growth. On the other hand there was a large difference between the levels observed in root and leaf samples from two-month-old plants; root signal levels were about the same as those of cell culture samples, whereas leaf levels were much reduced. These

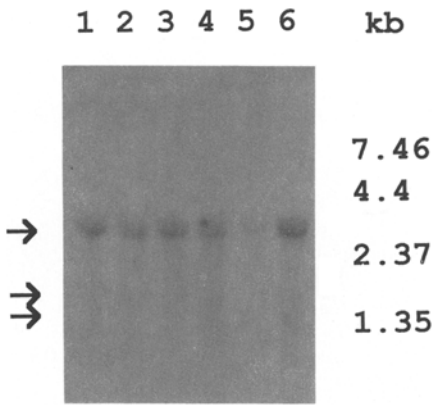


Fig. 3. Northern blot of carrot RNA probed with the AK-HSDH clone. Autoradiograph of RNA expression. Total RNA was isolated from carrot cell suspension cultures and 10 μg was loaded in each lane (lane 1, 1 day after transfer; lane 2, 3 days; lane 3, 5 days; lane 4, 7 days); 2-month old leaves (lane 5); and 2-month old roots (lane 6) of total RNA from each sample was separated on an agarose-formaldehyde-formamide gel and transferred to nitrocellulose membrane. The blot was hybridized with a ^{32}P-labelled 300 bp *Eco* RI-*Hinc* II fragment from the 5′ end of the clone.

mRNA levels corresponded with previously observed enzyme activity levels [14].

Genomic carrot DNA was also probed with the AK-HSDH clone (Fig. 4). DNA was digested with several restriction endonucleases. Blots were probed at low stringency and washed at both low and high temperatures. All the blots revealed fairly simple banding patterns, indicating a low number of hybridizing sequences in the genome. The large size of the fragments may indicate the presence of introns as are present in a soybean genomic clone (Gebhardt and Matthews, unpublished). Soybean genomic blots probed with the carrot HSDH clone showed a much more complex pattern (not shown). If there existed related, but only partially homologous, sequences in the carrot genome, higher wash temperatures should have resulted in the elimination of the bands for these sequences.

Fig. 2. Comparison of carrot AK-HSDH and *E. coli* AKI-HSDHI. The deduced amino acid sequence of the carrot HSDH was compared to the *E. coli* AKI-HSDHI using the GAP program of the University of Wisconsin Genetics Computer Group. Vertical lines indicate identical amino acids; dots indicate amino acids encoded by similar codons (double dot: codons differing in one nucleotide; single dots: codons differing in two nucleotides). An asterisk (*) under a pair of amino acids indicates identities shared by carrot AK-HSDH, the three *E. coli* proteins, *B. subtilis* AKII and yeast AK, a circle (o) indicates identities shared by the carrot protein, *E. coli* HSDHI and HSDHII, *B. subtilis* HSDH and *C. glutamicum* HSDH.

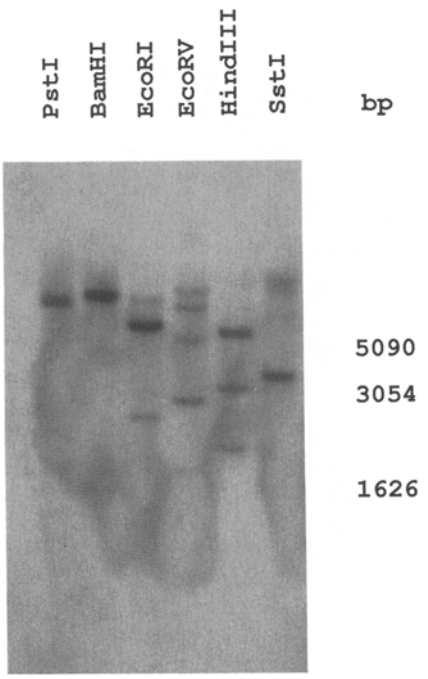


Fig. 4. Southern blot of carrot genomic DNA probed with the AK-HSDH cDNA clone. Carrot leaf DNA was digested with the indicated restriction enzymes, run on an agarose gel and blotted to nitrocellulose membrane. The blot was probed with the full-length AK-HSDH clone.

This was not the case, however, as higher wash temperatures simply reduced the intensity of all bands (not shown).

Discussion

We have isolated a cDNA clone from carrot that encodes a bifunctional AK-HSDH protein. Two lines of evidence for the bifunctional nature are (1) the strong amino acid homology of the carrot to the two bifunctional *E. coli* proteins, AKI-HSDHI and AKII-HSDHII; (2) the carrot protein (from which the PCR primers were designed) has both AK and HSDH function [24]. There have been no previous reports of a bifunctional AK-HSDH protein in plants. Furthermore, in yeast and in bacterial systems other than *E. coli*, the two enzymes have only been found to exist as separate proteins. Since this is the first AK or HSDH clone isolated from a plant, it is not known if such fused genes are common in plants. However, we have recently isolated a partial cDNA clone from soybean that has approximately 80% identity with the carrot clone and also appears to encode both AK and HSDH (unpublished).

Although the cDNA clone has a potential start site for translation, it is doubtful that this is the actual start. HSDH and AK have been shown to be localized in the chloroplasts of several plant species [2, 11, 15, 19, 23] and are likely to be located in the chloroplast of carrot. It follows that the carrot AK-HSDH probably has a chloroplast transit peptide at the amino terminus, since all known proteins transported into the chloroplast have been found to possess transit peptides [9]. The reading frame of the clone is continuous from the second base pair through the conserved KFGGT sequence which is found near the amino-terminus of all AK proteins. The 75–80 amino acids encoded by the 5′ end of the clone have the character of chloroplast transit peptides: the string is rich in serine, threonine, and small hydrophobic amino acids [9]. The sequence of the amino-terminus of the purified HSDH could not be determined and no precursor with a transit peptide has been identified. Further information is needed to resolve this question.

The deduced amino acid sequence of the carrot AK-HSDH shows homology to bacterial AK and HSDH proteins and to yeast AK. The most striking homology is to *E. coli* AKI-HSDHI. The overall homology between the carrot HSDH and *E. coli* AKI-HSDHI (48%) is higher than that between the two *E. coli* AK-HSDH proteins (46%) [9]. It has been proposed that AKI-HSDHI consists of three domains: an amino-terminal domain with AK activity (amino acids 1 to ca. 248), a carboxy-terminal domain with HSDH activity (amino acids 471 through 820), and an interface domain (amino acids 249–470) [4, 5]. The regions of the carrot protein that are most similar to AKI-HSDHI are the interface and the HSDH domains. In contrast, the homology between the carrot protein and *E. coli* AKII-HSDHII is lowest in the interface region. The homology in the other regions is only slightly lower

for AKII-HSDHII than for AKI-HSDHI. Both the carrot enzyme and AKI-HSDHI are regulated by threonine, whereas AKII-HSDHII is not [5, 14]. The higher homology between the two proteins may be related to this regulation.

In order to understand the regulation of these genes and the enzyme activities we need to know the number of genes and the number of distinct forms of the enzymes that exist. Studies on the HSDH enzyme of carrot indicate that there is only one HSDH protein, but this protein has different subunit structure and different regulatory properties *in vitro* depending on solution conditions [13]. Studies on the AK function have shown three forms of enzyme which differ in their sensitivity to lysine and threonine [24]. Some of these forms appear to interconvert. There is no conclusive evidence yet as to whether all the forms of carrot AK and HSDH are separate proteins or if they are forms of the same protein. DNA blots probed with the carrot cDNA indicated a low number of hybridizing sequences in the genome. RNA blots showed predominantly one band of about 3000 nucleotides. In contrast to the carrot results, experiments with a soybean clone show a much more complex pattern of DNA and RNA hybridization (data not shown). It is possible that carrot has only the one AK-HSDH which has both functions and is regulated both by threonine and lysine.

In most bacterial genera separate proteins are responsible for AK and HSDH functions [6, 18]. In these systems there is usually one AK which is regulated by concerted feedback by several end products, most often threonine and lysine. *E. coli* presents a different scheme in two ways: it has multifunctional AK-HSDH proteins and there is a separate AK for each end product (lysine, threonine, and methionine). Thus far we have identified a single multifunctional AK-HSDH in carrot. If this is the only AK or HSDH in carrot, it would represent a system that combines aspects of the various bacterial systems. It should be interesting to compare AK and HSDH genes from other plants and from blue-green algae to the known genes of carrot and bacteria.

Acknowledgement

This work was supported in part by the United States Department of Agriculture Office of Competitive Grant (No 87-CRCR-1-2284).

References

1. Bryan JK: Synthesis of the aspartate family and branched-chain amino acids. In: Miflin BJ (ed) The Biochemistry of Plants: A comprehensive Treatise, vol. 5, pp. 329–357. Academic Press, New York (1980).
2. Bryan JK, Lissik EA, Matthews BF: Changes in enzyme regulation during growth of maize. III. Intracellular localization of homoserine dehydrogenase in chloroplasts. Plant Physiol 59: 673–679 (1977).
3. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WI: Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294–5299 (1979).
4. Cohen GN: The common pathway to lysine, methionine and threonine. In: Herrmann KM, Somerville RL (eds) Amino Acids: Biosynthesis and Genetic Regulation, pp. 147–171. Addison-Wesley, Reading, MA (1983).
5. Cohen GN, Saint-Girons I: Biosynthesis of threonine, lysine and methionine. In: Neidhardt FC (ed) *Escherichia coli and Salmonella typhimurium:* Cellular and Molecular Biology, pp. 429–444. American Society for Microbiology, Washington, DC (1987).
6. Cohen GN, Stanier RY, Le Bras G: Regulation of the biosynthesis of the amino acids of the aspartate family in coliform bacteria and pseudomonads. J Bact 99: 791–801 (1969).
7. Feinberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
8. Kaneko T, Hashimoto T, Kumpaisal R, Yamada Y: Molecular cloning of wheat dihydrodipicolinate synthase. J Biol Chem 265: 17451–17455 (1990).
9. Keegstra K, Olsen LJ, Theg SM: Chloroplastic precursors and their transport across the envelope membranes. Annu Rev Plant Physiol Plant Mol Biol 40: 471–501 (1989).
10. Keim P, Olson TC, Shoemaker RC: A rapid protocol for isolating soybean DNA. Soybean Genet Newsl 15: 150–152 (1988).
11. Lea PJ, Mills WR, Miflin BJ: The isolation of a lysine-sensitive aspartate kinase from pea leaves and its involvement in homoserine biosynthesis in isolated chloroplasts. FEBS Lett 98: 165–168 (1979).
12. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
13. Matthews BF, Farrar MJ, Gray AC: Purification and in-

terconversion of homoserine dehydrogenase from *Daucus carota* cell suspension cultures. Plant Physiol 91: 1569–1574 (1989).
14. Matthews BF, Widholm JM: Regulation of lysine and threonine synthesis in carrot cell suspension cultures and whole carrot roots. Planta 141: 315–321 (1978).
15. Mills WR, Wilson KG: Effects of lysine, threonine and methionine on light-driven protein synthesis in isolated pea (*Pisum sativum* L.) chloroplasts. Planta 142: 153–160 (1978).
16. Parsot C, Cohen GN: Cloning and nucleotide sequence of the *Bacillus subtilis* hom gene coding for homoserine dehydrogenase. J Biol Chem 263: 14654–14660 (1988).
17. Rafalski JA, Falco SC: Structure of the yeast HOM gene which encodes aspartokinase. J Biol Chem 263: 2146–2151 (1988).
18. Robert-Gero M, Le Borgne L, Cohen GN: Concerted feedback inhibition of the aspartokinase of *Rhodospirrilum tenue* by threonine and methionine: a novel pattern. J Bact 112: 251–258 (1972).
19. Sainis JK, Mayne RG, Wallsgrove RM, Lea PJ, Miflin-BJ: Localization and characterization of homoserine dehydrogenase isolated from barley and pea leaves. Planta 152: 491–496 (1981).
20. Sanger F, Nicklen S, Coulson AR: DNA sequencing with the chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
21. Turano FJ, Jordan RL, Matthews BF: Immunological characterization of in vitro forms of homoserine dehydrogenase from carrot suspension cultures. Plant Physiol 92: 395–400 (1990).
22. Turano FJ, Weisemann JM, Matthews BF: Identification and expression of a cDNA clone encoding aspartate aminotransferase in carrot. Plant Physiol 100: 374–381 (1992).
23. Wallsgrove RM, Lea PJ, Miflin BJ: Intracellular localization of aspartate kinase and the enzymes of threonine and methionine biosynthesis in green leaves. Plant Physiol 71: 780–784 (1983).
24. Wilson BJ, Gray AC, Matthews BF: Bifunctional protein in carrot contains both aspartokinase and homoserine dehydrogenase activities. Plant Physiol 97: 1323–1328 (1991).
25. Zakin MM, Duchange N, Ferrara P, Cohen GN: Nucleotide sequence of the metL gene of *Escherichia coli*. J Biol Chem 258: 3028–3031 (1983).